

REMARKS

Applicants respectfully request entry of the following amendments and remarks in response to the Final Office Action mailed February 2, 2009. Applicants respectfully submit that the amendments and remarks contained herein place the instant application in condition for allowance.

Upon entry of the amendments in this response, claims 1, 6, 11 – 14, 16 – 17 and 19 – 39 are pending. In particular, Applicants amend claims 1, 6, 23 – 25, and 30 – 32. Reconsideration and allowance of the application and presently pending claims are respectfully requested.

I. Examiner Interview

Applicants first wish to express their sincere appreciation for the time that Examiner Macilwinen spent with Applicants' Attorney, Anthony Bonner, during a telephone discussion on March 25, 2009 regarding the outstanding Office Action. During that conversation, Examiner Macilwinen and Mr. Bonner discussed potential arguments and amendments with regard to claim 1, in view of the cited art. The general thrust of the potential principal arguments included a discussion of at least one embodiment of the present application disclosing "generating a phonetic equivalent for each word that includes only alphabetic displaying characters that have a phonetic equivalent." Thus, Applicants respectfully request that Examiner Macilwinen carefully consider this response and the amendments.

II. Rejections under 35 U.S.C. §103(a)

A. Claim 1 is Allowable Over *Shipp*, *Devine*, *Milliken*, *Anderson*, and *Uuencode* and MIME FAQ

The Office Action indicates that claim 1 stands rejected under 35 U.S.C. §103(a) as allegedly being unpatentable over U.S. Patent Publication Number 2004/0093384 ("*Shipp*") in view of U.S. Patent Number 6,968,571 ("*Devine*") further in view of U.S. Patent Publication Number 2004/0073617 ("*Milliken*") further in view of U.S. Patent Publication Number 2004/0064537 ("*Anderson*") further in view of Uuencode and MIME FAQ ("*Uuencode*"), further in view of U.S. Patent Number 6,732,157 ("*Gordon*"), further in view of A Bayesian Approach to Filtering Junk E-Mail ("*Sahami*") further in view of Identifying Junk Electronic Mail In Microsoft Outlook with a Support Vector Machine ("*Woitaszek*"). Applicants respectfully traverse this rejection for at least the reason that *Shipp* in view of *Devine*, further in view of *Milliken*, further in view of *Anderson*, further in view of *Uuencode*, further in view of *Gordon*, further in view of *Sahami* further in view of *Woitaszek* fails to disclose, teach, or suggest all of the elements of claim 1. More specifically, claim 1 recites:

A method comprising:

- (A) receiving an email message from a simple mail transfer protocol (SMTP) server, the email message comprising displaying characters and non-displaying characters, the email message further comprising:
 - (A1) a 32-bit string indicative of the length of the email message;
 - (A2) a text body;
 - (A3) an SMTP email address;
 - (A4) a domain name corresponding to the SMTP email address;
 - (A5) an attachment;
- (B) *searching for the non-displaying characters in the email;*
- (C) *removing the searched non-displaying characters;*
- (D) *determining non-alphabetic displaying characters in the email;*
- (E) *filtering the determined non-alphabetic displaying characters from the email;*
- (F) *generating a phonetic equivalent for each word that*

includes only alphabetic displaying characters that has a phonetic equivalent;

(G) ***tokenizing the phonetic equivalents in the text body to generate tokens representative of words in the text;***

(H) tokenizing the SMTP email address to generate a token representative of the SMTP email address;

(I) tokenizing the domain name to generate a token that is representative domain name;

(J) tokenizing the attachment to generate a token that is representative of the attachment, wherein tokenizing comprises:

(J1) generating a 128-bit MD5 hash of the attachment;

(J2) appending the 32-bit string to the generated MD5 hash to produce a 160-bit number; and

(J3) UUencoding the 160-bit number to generate the token representative of the attachment;

(K) determining a probability value for each of the generated tokens;

(L) sorting the generated tokens in accordance with the corresponding determined spam probability value to determine a predefined number of interesting tokens, the predefined number of interesting tokens being a subset of the generated tokens;

(M) selecting a the predefined number of interesting tokens, the interesting tokens being the generated tokens having the greatest non-neutral probability values;

(N) performing a Bayesian analysis on the selected interesting tokens to generate a spam probability; and

(O) categorizing the email message as a function of the generated spam probability.

(Emphasis added).

Applicants submit that claim 1, as amended, is allowable over the cited art for at least the reason that none of the references, taken alone or in combination, discloses, teaches, or suggests a "method comprising... ***searching for the non-displaying characters in the email... removing the searched non-displaying characters... determining non-alphabetic displaying characters in the email... filtering the determined non-alphabetic displaying characters from the email... generating a phonetic equivalent for each word that includes only alphabetic displaying characters that has a phonetic equivalent... [and] tokenizing the phonetic equivalents in the text body to generate tokens representative of words in the text***" as recited in claim 1, as amended. While the Office Action argues that a combination of eight (8) references renders claim 1 obvious, Applicants disagree. More specifically, *Shipp*

discloses "[t]he invention is to weed out candidates for logging so that the normal mail is not logged. This reduces the burden on the database 23, and improves performance... A simplistic algorithm would be: If mail contains attachments, do not log" (page 3, paragraphs [0080] – [0081]). However, *Shipp* fails to even suggest "***searching for the non-displaying characters in the email... removing the searched non-displaying characters... determining non-alphabetic displaying characters in the email... filtering the determined non-alphabetic displaying characters from the email... generating a phonetic equivalent for each word that includes only alphabetic displaying characters that has a phonetic equivalent... [and] tokenizing the phonetic equivalents in the text body to generate tokens representative of words in the text***" as recited in claim 1, as amended.

Additionally, *Devine* fails to overcome the deficiencies of *Shipp*. More specifically, *Devine* discloses a "series of security protocols and an integrated system for the same that enables a user to interface with one or more application services provided by remote servers over the public internet, or an enterprise extranet" (column 2, line 58). However, *Devine* fails to even suggest "***searching for the non-displaying characters in the email... removing the searched non-displaying characters... determining non-alphabetic displaying characters in the email... filtering the determined non-alphabetic displaying characters from the email... generating a phonetic equivalent for each word that includes only alphabetic displaying characters that has a phonetic equivalent... [and] tokenizing the phonetic equivalents in the text body to generate tokens representative of words in the text***" as recited in claim 1, as amended.

Further, *Milliken* fails to overcome the deficiencies of *Shipp* and *Devine*. More specifically, *Milliken* discloses a "method for detecting transmission of potentially unwanted e-mail messages is provided. The method includes receiving e-mail messages and generating hash values based on one or more portions of the e-mail messages" (paragraph [0010]).

However, *Milliken* fails to even suggest ***"searching for the non-displaying characters in the email... removing the searched non-displaying characters... determining non-alphabetic displaying characters in the email... filtering the determined non-alphabetic displaying characters from the email... generating a phonetic equivalent for each word that includes only alphabetic displaying characters that has a phonetic equivalent... [and] tokenizing the phonetic equivalents in the text body to generate tokens representative of words in the text"*** as recited in claim 1, as amended.

Similarly, *Anderson* fails to overcome the deficiencies of *Shipp*, *Devine*, and *Milliken*. More specifically, *Anderson* discloses "data payloads containing the identified network transmission items are selectively transmitted on an internal destination node within an internal network" (page 7, paragraph [0076]). However, *Anderson* fails to even suggest ***"searching for the non-displaying characters in the email... removing the searched non-displaying characters... determining non-alphabetic displaying characters in the email... filtering the determined non-alphabetic displaying characters from the email... generating a phonetic equivalent for each word that includes only alphabetic displaying characters that has a phonetic equivalent... [and] tokenizing the phonetic equivalents in the text body to generate tokens representative of words in the text"*** as recited in claim 1, as amended.

Additionally, *Uuencode* fails to overcome the deficiencies of *Shipp*, *Devine*, *Milliken*, and *Anderson*. More specifically, *Uuencode* discloses "discloses converting a binary file on an ASCII or text file so it can be sent as an attachment to an email message or downloaded from a newsgroup" (page 1, line 14). However, *Uuencode* fails to even suggest ***"searching for the non-displaying characters in the email... removing the searched non-displaying characters... determining non-alphabetic displaying characters in the email... filtering the determined non-alphabetic displaying characters from the email... generating a phonetic***

equivalent for each word that includes only alphabetic displaying characters that has a phonetic equivalent... [and] tokenizing the phonetic equivalents in the text body to generate tokens representative of words in the text' as recited in claim 1, as amended.

Further, *Gordon* fails to overcome the deficiencies of *Shipp*, *Devine*, *Milliken*, *Anderson*, and *Uencode*. More specifically, *Gordon* discloses "a probability associated with each of the words and/or groups of words [in an email message] is determined using the Bayes rules database" (emphasis added, column, 11, line 27). However, *Gordon* fails to even suggest ***"searching for the non-displaying characters in the email... removing the searched non-displaying characters... determining non-alphabetic displaying characters in the email... filtering the determined non-alphabetic displaying characters from the email... generating a phonetic equivalent for each word that includes only alphabetic displaying characters that has a phonetic equivalent... [and] tokenizing the phonetic equivalents in the text body to generate tokens representative of words in the text"*** as recited in claim 1, as amended.

Additionally, *Sahami* fails to overcome the deficiencies of *Shipp*, *Devine*, *Milliken*, *Anderson*, *Uencode*, and *Gordon*. More specifically, *Sahami* discloses "Sahami discloses "[determining] whether a message has attached documents (most junk E-mail does not have them)... [is] also [a] powerful distinguisher between junk and legitimate E-mail" (page 3, right column, last paragraph). However, *Sahami* fails to even suggest ***"searching for the non-displaying characters in the email... removing the searched non-displaying characters... determining non-alphabetic displaying characters in the email... filtering the determined non-alphabetic displaying characters from the email... generating a phonetic equivalent for each word that includes only alphabetic displaying characters that has a phonetic equivalent... [and] tokenizing the phonetic equivalents in the text body to generate tokens representative of words in the text"*** as recited in claim 1, as amended.

Further, *Woitaszek* fails to overcome the deficiencies of *Shipp*, *Devine*, *Milliken*, *Anderson*, *Uencode*, *Gordon*, and *Sahami*. More specifically, *Woitaszek* discloses "[e]ach message [is] parsed to completely remove any headers, attachments, HTML markup, punctuation, and extended characters. This procedure essentially reduces a mail message to a series of delimited lowercase string tokens" (page 2, section 4). However, *Woitaszek* fails to even suggest "***searching for the non-displaying characters in the email... removing the searched non-displaying characters... determining non-alphabetic displaying characters in the email... filtering the determined non-alphabetic displaying characters from the email... generating a phonetic equivalent for each word that includes only alphabetic displaying characters that has a phonetic equivalent... [and] tokenizing the phonetic equivalents in the text body to generate tokens representative of words in the text***" as recited in claim 1, as amended. Thus, for at least the reason that none of the references disclose or suggest all of the claimed elements, the combination of eight (8) references does not render obvious claim 1. For at least these reasons, claim 1, as amended, is allowable.

B. Claim 39 is Allowable Over *Shipp*, *Devine*, *Milliken*, *Anderson*, and *Uencode* and MIME FAQ

The Office Action indicates that claim 39 stands rejected under 35 U.S.C. §103(a) as allegedly being unpatentable over U.S. Patent Number 2004/0093384 ("*Shipp*") in view of U.S. Patent Number 6,968,571 ("*Devine*"), further in view of U.S. Publication Number 2004/0073617 ("*Milliken*") further in view of U.S. Publication Number 2004/0064537 ("*Anderson*") further in view of *Uencode* and *MIME FAQ*. Applicants respectfully traverse this rejection for at least the reason that *Shipp* in view of *Devine*, further in view of *Milliken*, further in view of *Anderson*, further in view of *Uencode* and *MIME FAQ* fail to disclose, teach, or suggest all of the elements of claim 39. More specifically, dependent claim 39 is believed to be allowable for at least the

reason that this claim depends from and includes the elements of allowable independent claim

1. *In re Fine, Minnesota Mining and Mfg.Co. v. Chemque, Inc.*, 303 F.3d 1294, 1299 (Fed. Cir. 2002).

C. Claim 6 is Allowable Over Shipp, Milliken, Sahami, and Woitaszek

The Office Action indicates that claim 6 stands rejected under 35 U.S.C. §103(a) as allegedly being unpatentable over U.S. Patent Number 2004/0093384 ("*Shipp*") in view of U.S. Publication Number 2004/0073617 ("*Milliken*") further in view of A Bayesian Approach to Filtering Junk E-Mail ("*Sahami*") further in view of Identifying Junk Electronic Mail In Microsoft Outlook with a Support Vector Machine ("*Woitaszek*"). Applicants respectfully traverse this rejection for at least the reason that *Shipp* further in view of *Milliken*, further in view of *Sahami*, further in view of *Woitaszek* fail to disclose, teach, or suggest all of the elements of claim 6. More specifically, claim 6 recites:

A method comprising:

receiving an email message comprising a text body, an SMTP email address, an attachment, and a domain name corresponding to the SMTP email address, the text body including displaying characters and non-displaying characters;

searching for the non-displaying characters in the email;

removing the searched non-displaying characters;

tokenizing the SMTP email address to generate a token representative of the displaying characters of the SMTP email address;

tokenizing the attachment to generate a token that is representative of the attachment;

tokenizing the domain name to generate a token representative of the domain name;

determining a spam probability value from the generated tokens; and

sorting the generated tokens in accordance with the corresponding determined spam probability value to determine a predefined number of interesting tokens, the predefined number of interesting tokens being a subset of the generated tokens.

(Emphasis added).

Applicants submit that claim 6, as amended, is allowable over the cited art for at least the reason that none of the references, taken alone or in combination, discloses, teaches, or suggests a "method comprising... **searching for the non-displaying characters in the email... removing the searched non-displaying characters...** [and] **tokenizing the SMTP email address to generate a token representative of the displaying characters of the SMTP email address**" as recited in claim 6, as amended. While the Office Action argues that a combination of references renders claim 6 obvious, Applicants disagree. More specifically, *Shipp* discloses "[t]he invention is to weed out candidates for logging so that the normal mail is not logged. This reduces the burden on the database 23, and improves performance... A simplistic algorithm would be: If mail contains attachments, do not log" (page 3, paragraphs [0080] – [0081]). However, *Shipp* fails to even suggest "**searching for the non-displaying characters in the email... removing the searched non-displaying characters...** [and] **tokenizing the SMTP email address to generate a token representative of the displaying characters of the SMTP email address**" as recited in claim 6, as amended.

Further, *Milliken* fails to overcome the deficiencies of *Shipp*. More specifically, *Milliken* discloses a "method for detecting transmission of potentially unwanted e-mail messages is provided. The method includes receiving e-mail messages and generating hash values based on one or more portions of the e-mail messages" (paragraph [0010]). However, *Milliken* fails to even suggest "**searching for the non-displaying characters in the email... removing the searched non-displaying characters...** [and] **tokenizing the SMTP email address to generate a token representative of the displaying characters of the SMTP email address**" as recited in claim 6, as amended.

Additionally, *Sahami* fails to overcome the deficiencies of *Shipp*, and *Milliken*. More specifically, *Sahami* discloses "*Sahami* discloses "[determining] whether a message has attached documents (most junk E-mail does not have them)... [is] also [a] powerful distinguisher

between junk and legitimate E-mail" (page 3, right column, last paragraph). However, *Sahami* fails to even suggest "***searching for the non-displaying characters in the email... removing the searched non-displaying characters... [and] tokenizing the SMTP email address to generate a token representative of the displaying characters of the SMTP email address***" as recited in claim 6, as amended.

Further, *Woitaszek* fails to overcome the deficiencies of *Shipp*, *Milliken*, and *Sahami*. More specifically, *Woitaszek* discloses "[e]ach message [is] parsed to completely remove any headers, attachments, HTML markup, punctuation, and extended characters. This procedure essentially reduces a mail message to a series of delimited lowercase string tokens" (page 2, section 4). However, *Woitaszek* fails to even suggest "***searching for the non-displaying characters in the email... removing the searched non-displaying characters... [and] tokenizing the SMTP email address to generate a token representative of the displaying characters of the SMTP email address***" as recited in claim 6, as amended. Thus, for at least the reason that none of the references disclose or suggest all of the claimed elements, the combination of references does not render obvious claim 6. For at least these reasons, claim 6, as amended, is allowable.

D. Claims 11 – 14, 16 – 17, and 19 – 22 are Allowable Over *Shipp*, *Milliken*, *Sahami*, and *Woitaszek*

The Office Action indicates that claims 11 – 14, 16 – 17, and 19 – 22 stand rejected under 35 U.S.C. §103(a) as allegedly being unpatentable over U.S. Patent Number 2004/0093384 ("*Shipp*") in view of U.S. Publication Number 2004/0073617 ("*Milliken*") further in view of A Bayesian Approach to Filtering Junk E-Mail ("*Sahami*") further in view of Identifying Junk Electronic Mail In Microsoft Outlook with a Support Vector Machine ("*Woitaszek*"). Applicants respectfully traverse this rejection for at least the reason that *Shipp* further in view of

Milliken, further in view of *Sahami*, further in view of *Woitaszek* fail to disclose, teach, or suggest all of the elements of claims 11 – 14, 16 – 17, and 19 – 22. More specifically, dependent claims 11 – 14, 16 – 17, and 19 – 22 are believed to be allowable for at least the reason that these claims depend from and include the elements of allowable independent claim 6. *In re Fine*, *Minnesota Mining and Mfg. Co. v. Chemque, Inc.*, 303 F.3d 1294, 1299 (Fed. Cir. 2002).

E. Claim 23 is Allowable Over *Shipp*, *Milliken*, *Sahami*, and *Woitaszek*

The Office Action indicates that claim 23 stands rejected under 35 U.S.C. §103(a) as allegedly being unpatentable over U.S. Patent Number 2004/0093384 ("*Shipp*") in view of U.S. Publication Number 2004/0073617 ("*Milliken*") further in view of A Bayesian Approach to Filtering Junk E-Mail ("*Sahami*") further in view of Identifying Junk Electronic Mail In Microsoft Outlook with a Support Vector Machine ("*Woitaszek*"). Applicants respectfully traverse this rejection for at least the reason that *Shipp* further in view of *Milliken*, further in view of *Sahami*, further in view of *Woitaszek* fail to disclose, teach, or suggest all of the elements of claim 23. More specifically, claim 23 recites:

A system comprising:
email receive logic configured to receive an email message comprising an SMTP email address, a domain name corresponding to the SMTP email address, and an address; address, the email message further including displaying characters and non-displaying characters;
searching logic configured to search for the non-displaying characters in the email;
removing logic configured to remove the searched non-displaying characters;
tokenize logic configured to tokenize the SMTP email address to generate a token representative of the SMTP email address;
tokenize logic configured to tokenize the attachment to generate a token that is representative of the attachment;
tokenize logic configured to tokenize the domain name to generate a token representative of the domain name;
analysis logic configured to determine a spam probability value from the generated tokens; and

sorting logic configured to sort the generated tokens in accordance with the corresponding determined spam probability value to determine a predefined number of interesting tokens, the predefined number of interesting tokens being a subset of the generated tokens, **wherein only displaying characters are tokenized.**

(Emphasis added).

Applicants submit that claim 23, as amended, is allowable over the cited art for at least the reason that none of the references, taken alone or in combination, discloses, teaches, or suggests a "system comprising... **searching logic configured to search for the non-displaying characters in the email... [and] removing logic configured to remove the searched non-displaying characters... wherein only displaying characters are tokenized**" as recited in claim 23, as amended. While the Office Action argues that a combination of references renders claim 23 obvious, Applicants disagree. More specifically, *Shipp* discloses "[t]he invention is to weed out candidates for logging so that the normal mail is not logged. This reduces the burden on the database 23, and improves performance... A simplistic algorithm would be: If mail contains attachments, do not log" (page 3, paragraphs [0080] – [0081]). However, *Shipp* fails to even suggest "**searching logic configured to search for the non-displaying characters in the email... [and] removing logic configured to remove the searched non-displaying characters... wherein only displaying characters are tokenized**" as recited in claim 23, as amended.

Further, *Milliken* fails to overcome the deficiencies of *Shipp*. More specifically, *Milliken* discloses a "method for detecting transmission of potentially unwanted e-mail messages is provided. The method includes receiving e-mail messages and generating hash values based on one or more portions of the e-mail messages" (paragraph [0010]). However, *Milliken* fails to even suggest "**searching logic configured to search for the non-displaying characters in the email... [and] removing logic configured to remove the searched non-displaying**

characters... wherein only displaying characters are tokenized" as recited in claim 23, as amended.

Additionally, *Sahami* fails to overcome the deficiencies of *Shipp* and *Milliken*. More specifically, *Sahami* discloses "*Sahami* discloses "[determining] whether a message has attached documents (most junk E-mail does not have them)... [is] also [a] powerful distinguisher between junk and legitimate E-mail" (page 3, right column, last paragraph). However, *Sahami* fails to even suggest "**searching logic configured to search for the non-displaying characters in the email... [and] removing logic configured to remove the searched non-displaying characters... wherein only displaying characters are tokenized**" as recited in claim 23, as amended.

Further, *Woitaszek* fails to overcome the deficiencies of *Shipp*, *Milliken*, and *Sahami*. More specifically, *Woitaszek* discloses "[e]ach message [is] parsed to completely remove any headers, attachments, HTML markup, punctuation, and extended characters. This procedure essentially reduces a mail message to a series of delimited lowercase string tokens" (page 2, section 4). However, *Woitaszek* fails to even suggest "**searching logic configured to search for the non-displaying characters in the email... [and] removing logic configured to remove the searched non-displaying characters... wherein only displaying characters are tokenized**" as recited in claim 23, as amended. Thus, for at least the reason that none of the references disclose or suggest all of the claimed elements, the combination of references does not render obvious claim 23. For at least these reasons, claim 23, as amended, is allowable.

F. Claim 24 is Allowable Over *Shipp*, *Milliken*, *Sahami*, and *Woitaszek*

The Office Action indicates that claim 24 stands rejected under 35 U.S.C. §103(a) as allegedly being unpatentable over U.S. Patent Number 2004/0093384 ("*Shipp*") in view of U.S. Publication Number 2004/0073617 ("*Milliken*") further in view of A Bayesian Approach to

Filtering Junk E-Mail ("*Sahami*") further in view of Identifying Junk Electronic Mail In Microsoft Outlook with a Support Vector Machine ("*Woitaszek*"). Applicants respectfully traverse this rejection for at least the reason that *Shipp* further in view of *Milliken*, further in view of *Sahami*, further in view of *Woitaszek* fail to disclose, teach, or suggest all of the elements of claim 24. More specifically, claim 24 recites:

A system comprising:
means for receiving an email message comprising an SMTP email address, a domain name corresponding to the SMTP email address, and an address, the email message further including displaying characters and non-displaying characters;
means for searching for the non-displaying characters in the email;
means for removing the searched non-displaying characters;
means for tokenizing the SMTP email address to generate a token representative of the SMTP email address;
means for tokenizing the attachment to generate a token that is representative of the attachment;
means for tokenizing the domain name to generate a token representative of the domain name;
means for determining a spam probability value from the generated tokens; and
means for sorting the generated tokens in accordance with the corresponding determined spam probability value to determine a predefined number of interesting tokens, the predefined number of interesting tokens being a subset of the generated tokens,
wherein only displaying characters are tokenized.
(Emphasis added).

Applicants submit that claim 24, as amended, is allowable over the cited art for at least the reason that none of the references, taken alone or in combination, discloses, teaches, or suggests a "system comprising... ***means for searching for the non-displaying characters in the email... [and] means for removing the searched non-displaying characters... wherein only displaying characters are tokenized***" as recited in claim 24, as amended. While the Office Action argues that a combination of references renders claim 24 obvious, Applicants disagree. More specifically, *Shipp* discloses "[t]he invention is to weed out candidates for logging so that the normal mail is not logged. This reduces the burden on the database 23, and

improves performance... A simplistic algorithm would be: If mail contains attachments, do not log" (page 3, paragraphs [0080] – [0081]). However, *Shipp* fails to even suggest "***means for searching for the non-displaying characters in the email... [and] means for removing the searched non-displaying characters... wherein only displaying characters are tokenized***" as recited in claim 24, as amended.

Further, *Milliken* fails to overcome the deficiencies of *Shipp*. More specifically, *Milliken* discloses a "method for detecting transmission of potentially unwanted e-mail messages is provided. The method includes receiving e-mail messages and generating hash values based on one or more portions of the e-mail messages" (paragraph [0010]). However, *Milliken* fails to even suggest "***means for searching for the non-displaying characters in the email... [and] means for removing the searched non-displaying characters... wherein only displaying characters are tokenized***" as recited in claim 24, as amended.

Additionally, *Sahami* fails to overcome the deficiencies of *Shipp* and *Milliken*. More specifically, *Sahami* discloses "*Sahami* discloses "[determining] whether a message has attached documents (most junk E-mail does not have them)... [is] also [a] powerful distinguisher between junk and legitimate E-mail" (page 3, right column, last paragraph). However, *Sahami* fails to even suggest "***means for searching for the non-displaying characters in the email... [and] means for removing the searched non-displaying characters... wherein only displaying characters are tokenized***" as recited in claim 24, as amended.

Further, *Woitaszek* fails to overcome the deficiencies of *Shipp*, *Milliken*, and *Sahami*. More specifically, *Woitaszek* discloses "[e]ach message [is] parsed to completely remove any headers, attachments, HTML markup, punctuation, and extended characters. This procedure essentially reduces a mail message to a series of delimited lowercase string tokens" (page 2, section 4). However, *Woitaszek* fails to even suggest "***means for searching for the non-displaying characters in the email... [and] means for removing the searched non-***

displaying characters... wherein only displaying characters are tokenized" as recited in claim 24, as amended. Thus, for at least the reason that none of the references disclose or suggest all of the claimed elements, the combination of references does not render obvious claim 24. For at least these reasons, claim 24, as amended, is allowable.

G. Claim 25 is Allowable Over Shipp, Milliken, Sahami, and Woitaszek

The Office Action indicates that claim 25 stands rejected under 35 U.S.C. §103(a) as allegedly being unpatentable over U.S. Patent Number 2004/0093384 ("*Shipp*") in view of U.S. Publication Number 2004/0073617 ("*Milliken*") further in view of A Bayesian Approach to Filtering Junk E-Mail ("*Sahami*") further in view of Identifying Junk Electronic Mail In Microsoft Outlook with a Support Vector Machine ("*Woitaszek*"). Applicants respectfully traverse this rejection for at least the reason that *Shipp* further in view of *Milliken*, further in view of *Sahami*, further in view of *Woitaszek* fail to disclose, teach, or suggest all of the elements of claim 25. More specifically, claim 25 recites:

A computer-readable medium that includes a program that, when executed by a computer, performs at least the following:

receive an email message comprising an SMTP email address, a domain name corresponding to the SMTP email address, and an attachment, the email message further including displaying characters and non-displaying characters;

search for the non-displaying characters in the email;

remove the searched non-displaying characters;

tokenize the SMTP email address to generate a token representative of the SMTP email address;

tokenize the attachment to generate a token that is representative of the attachment;

tokenize the domain name to generate a token representative of the domain name;

determine a spam probability value from the generated tokens; and

sort the generated tokens in accordance with the corresponding determined spam probability value to determine a predefined number of interesting tokens, the predefined number of interesting tokens being a subset of the generated tokens, ***wherein only displaying characters are tokenized.***

(Emphasis added).

Applicants submit that claim 25, as amended, is allowable over the cited art for at least the reason that none of the references, taken alone or in combination, discloses, teaches, or suggests a "computer-readable medium that includes a program that, when executed by a computer, performs at least the following... ***search for the non-displaying characters in the email... [and] remove the searched non-displaying characters... wherein only displaying characters are tokenized***" as recited in claim 25, as amended. While the Office Action argues that a combination of references renders claim 25 obvious, Applicants disagree. More specifically, *Shipp* discloses "[t]he invention is to weed out candidates for logging so that the normal mail is not logged. This reduces the burden on the database 23, and improves performance... A simplistic algorithm would be: If mail contains attachments, do not log" (page 3, paragraphs [0080] – [0081]). However, *Shipp* fails to even suggest "computer-readable medium that includes a program that, when executed by a computer, performs at least the following... ***search for the non-displaying characters in the email... [and] remove the searched non-displaying characters... wherein only displaying characters are tokenized***" as recited in claim 25, as amended.

Further, *Milliken* fails to overcome the deficiencies of *Shipp*. More specifically, *Milliken* discloses a "method for detecting transmission of potentially unwanted e-mail messages is provided. The method includes receiving e-mail messages and generating hash values based on one or more portions of the e-mail messages" (paragraph [0010]). However, *Milliken* fails to even suggest "computer-readable medium that includes a program that, when executed by a computer, performs at least the following... ***search for the non-displaying characters in the email... [and] remove the searched non-displaying characters... wherein only displaying characters are tokenized***" as recited in claim 25, as amended.

Additionally, *Sahami* fails to overcome the deficiencies of *Shipp* and *Milliken*. More specifically, *Sahami* discloses "*Sahami* discloses "[determining] whether a message has attached documents (most junk E-mail does not have them)... [is] also [a] powerful distinguisher between junk and legitimate E-mail" (page 3, right column, last paragraph). However, *Sahami* fails to even suggest "computer-readable medium that includes a program that, when executed by a computer, performs at least the following... **search for the non-displaying characters in the email...** [and] **remove the searched non-displaying characters... wherein only displaying characters are tokenized**" as recited in claim 25, as amended.

Further, *Woitaszek* fails to overcome the deficiencies of *Shipp*, *Milliken*, and *Sahami*. More specifically, *Woitaszek* discloses "[e]ach message [is] parsed to completely remove any headers, attachments, HTML markup, punctuation, and extended characters. This procedure essentially reduces a mail message to a series of delimited lowercase string tokens" (page 2, section 4). However, *Woitaszek* fails to even suggest "computer-readable medium that includes a program that, when executed by a computer, performs at least the following... **search for the non-displaying characters in the email...** [and] **remove the searched non-displaying characters... wherein only displaying characters are tokenized**" as recited in claim 25, as amended. Thus, for at least the reason that none of the references disclose or suggest all of the claimed elements, the combination of references does not render obvious claim 25. For at least these reasons, claim 25, as amended, is allowable.

H. Claims 26 – 29 are Allowable Over *Shipp*, *Milliken*, *Sahami*, and *Woitaszek*

The Office Action indicates that claims 26 – 29 stand rejected under 35 U.S.C. §103(a) as allegedly being unpatentable over U.S. Patent Number 2004/0093384 ("*Shipp*") in view of U.S. Publication Number 2004/0073617 ("*Milliken*") further in view of A Bayesian Approach to Filtering Junk E-Mail ("*Sahami*") further in view of Identifying Junk Electronic Mail In Microsoft

Outlook with a Support Vector Machine ("*Woitaszek*"). Applicants respectfully traverse this rejection for at least the reason that *Shipp* further in view of *Milliken*, further in view of *Sahami*, further in view of *Woitaszek* fail to disclose, teach, or suggest all of the elements of claims 25 – 29. More specifically, dependent claims 26 – 29 are believed to be allowable for at least the reason that these claims depend from and include the elements of allowable independent claim 25. *In re Fine, Minnesota Mining and Mfg. Co. v. Chemque, Inc.*, 303 F.3d 1294, 1299 (Fed. Cir. 2002).

I. Claim 30 is Allowable Over *Shipp*, *Milliken*, *Sahami*, and *Woitaszek*

The Office Action indicates that claim 30 stands rejected under 35 U.S.C. §103(a) as allegedly being unpatentable over U.S. Patent Number 2004/0093384 ("*Shipp*") in view of U.S. Publication Number 2004/0073617 ("*Milliken*") further in view of A Bayesian Approach to Filtering Junk E-Mail ("*Sahami*") further in view of Identifying Junk Electronic Mail In Microsoft Outlook with a Support Vector Machine ("*Woitaszek*"). Applicants respectfully traverse this rejection for at least the reason that *Shipp* further in view of *Milliken*, further in view of *Sahami*, further in view of *Woitaszek* fail to disclose, teach, or suggest all of the elements of claim 30. More specifically, claim 30 recites:

A system comprising:
a memory component that stores at least the following:
email receive logic configured to receive an email message comprising an address, the email message further including displaying characters and non-displaying characters;
search logic configured to search for the non-displaying characters in the email;
remove logic configured to remove the searched non-displaying characters;
tokenize logic configured to tokenize the attachment to generate a token representative of the attachment;
analysis logic configured to determine a spam probability value from the generated token; and
sort logic configured to sort the generated tokens in accordance with the corresponding spam probability value to

determine a predefined number of interesting tokens, the predefined number of interesting tokens being a subset of the generated tokens, **wherein only displaying characters are tokenized.**

(Emphasis added).

Applicants submit that claim 30, as amended, is allowable over the cited art for at least the reason that none of the references, taken alone or in combination, discloses, teaches, or suggests a "system comprising... **search logic configured to search for the non-displaying characters in the email...** [and] **remove logic configured to remove the searched non-displaying characters... wherein only displaying characters are tokenized**" as recited in claim 30, as amended. While the Office Action argues that a combination of references renders claim 30 obvious, Applicants disagree. More specifically, *Shipp* discloses "[t]he invention is to weed out candidates for logging so that the normal mail is not logged. This reduces the burden on the database 23, and improves performance... A simplistic algorithm would be: If mail contains attachments, do not log" (page 3, paragraphs [0080] – [0081]). However, *Shipp* fails to even suggest "**search logic configured to search for the non-displaying characters in the email...** [and] **remove logic configured to remove the searched non-displaying characters... wherein only displaying characters are tokenized**" as recited in claim 30, as amended.

Further, *Milliken* fails to overcome the deficiencies of *Shipp*. More specifically, *Milliken* discloses a "method for detecting transmission of potentially unwanted e-mail messages is provided. The method includes receiving e-mail messages and generating hash values based on one or more portions of the e-mail messages" (paragraph [0010]). However, *Milliken* fails to even suggest "**search logic configured to search for the non-displaying characters in the email...** [and] **remove logic configured to remove the searched non-displaying characters... wherein only displaying characters are tokenized**" as recited in claim 30, as amended.

Additionally, *Sahami* fails to overcome the deficiencies of *Shipp* and *Milliken*. More specifically, *Sahami* discloses "*Sahami* discloses "[determining] whether a message has attached documents (most junk E-mail does not have them)... [is] also [a] powerful distinguisher between junk and legitimate E-mail" (page 3, right column, last paragraph). However, *Sahami* fails to even suggest "***search logic configured to search for the non-displaying characters in the email... [and] remove logic configured to remove the searched non-displaying characters... wherein only displaying characters are tokenized***" as recited in claim 30, as amended.

Further, *Woitaszek* fails to overcome the deficiencies of *Shipp*, *Milliken*, and *Sahami*. More specifically, *Woitaszek* discloses "[e]ach message [is] parsed to completely remove any headers, attachments, HTML markup, punctuation, and extended characters. This procedure essentially reduces a mail message to a series of delimited lowercase string tokens" (page 2, section 4). However, *Woitaszek* fails to even suggest "***search logic configured to search for the non-displaying characters in the email... [and] remove logic configured to remove the searched non-displaying characters... wherein only displaying characters are tokenized***" as recited in claim 30, as amended. Thus, for at least the reason that none of the references disclose or suggest all of the claimed elements, the combination of references does not render obvious claim 30. For at least these reasons, claim 30, as amended, is allowable.

J. Claim 31 is Allowable Over *Shipp*, *Milliken*, *Sahami*, and *Woitaszek*

The Office Action indicates that claim 31 stands rejected under 35 U.S.C. §103(a) as allegedly being unpatentable over U.S. Patent Number 2004/0093384 ("*Shipp*") in view of U.S. Publication Number 2004/0073617 ("*Milliken*") further in view of A Bayesian Approach to Filtering Junk E-Mail ("*Sahami*") further in view of Identifying Junk Electronic Mail In Microsoft Outlook with a Support Vector Machine ("*Woitaszek*"). Applicants respectfully traverse this

rejection for at least the reason that *Shipp* further in view of *Milliken*, further in view of *Sahami*, further in view of *Woitaszek* fail to disclose, teach, or suggest all of the elements of claim 31.

More specifically, claim 31 recites:

A system comprising:
means for receiving an email message comprising an address, the email message further including displaying characters and non-displaying characters;
means for searching for the non-displaying characters in the email;
means for removing the searched non-displaying characters;
means for tokenizing the attachment to generate a token representative of the attachment;
means for determining a spam probability value from the generated token; and
means for sorting the generated tokens in accordance with the corresponding determined spam probability value to determine a predefined number of interesting tokens, the predefined number of interesting tokens being a subset of the generated tokens,
wherein only displaying characters are tokenized.
(Emphasis added).

Applicants submit that claim 31, as amended, is allowable over the cited art for at least the reason that none of the references, taken alone or in combination, discloses, teaches, or suggests a “system comprising... ***means for removing the searched non-displaying characters...*** [and] ***means for tokenizing the attachment to generate a token representative of the attachment... wherein only displaying characters are tokenized***” as recited in claim 31, as amended. While the Office Action argues that a combination of references renders claim 31 obvious, Applicants disagree. More specifically, *Shipp* discloses “[t]he invention is to weed out candidates for logging so that the normal mail is not logged. This reduces the burden on the database 23, and improves performance... A simplistic algorithm would be: If mail contains attachments, do not log” (page 3, paragraphs [0080] – [0081]). However, *Shipp* fails to even suggest “***means for removing the searched non-displaying characters...*** [and] ***means for tokenizing the attachment to generate a token***

representative of the attachment... wherein only displaying characters are tokenized as recited in claim 31, as amended.

Further, *Milliken* fails to overcome the deficiencies of *Shipp*. More specifically, *Milliken* discloses a "method for detecting transmission of potentially unwanted e-mail messages is provided. The method includes receiving e-mail messages and generating hash values based on one or more portions of the e-mail messages" (paragraph [0010]). However, *Milliken* fails to even suggest "***means for removing the searched non-displaying characters... [and] means for tokenizing the attachment to generate a token representative of the attachment... wherein only displaying characters are tokenized***" as recited in claim 31, as amended.

Additionally, *Sahami* fails to overcome the deficiencies of *Shipp* and *Milliken*. More specifically, *Sahami* discloses "*Sahami* discloses "[determining] whether a message has attached documents (most junk E-mail does not have them)... [is] also [a] powerful distinguisher between junk and legitimate E-mail" (page 3, right column, last paragraph). However, *Sahami* fails to even suggest "***means for removing the searched non-displaying characters... [and] means for tokenizing the attachment to generate a token representative of the attachment... wherein only displaying characters are tokenized***" as recited in claim 31, as amended.

Further, *Woitaszek* fails to overcome the deficiencies of *Shipp*, *Milliken*, and *Sahami*. More specifically, *Woitaszek* discloses "[e]ach message [is] parsed to completely remove any headers, attachments, HTML markup, punctuation, and extended characters. This procedure essentially reduces a mail message to a series of delimited lowercase string tokens" (page 2, section 4). However, *Woitaszek* fails to even suggest "***means for removing the searched non-displaying characters... [and] means for tokenizing the attachment to generate a token representative of the attachment... wherein only displaying characters are tokenized***" as recited in claim 31, as amended. Thus, for at least the reason that none of the

references disclose or suggest all of the claimed elements, the combination of references does not render obvious claim 31. For at least these reasons, claim 31, as amended, is allowable.

K. Claim 32 is Allowable Over Shipp, Milliken, Sahami, and Woitaszek

The Office Action indicates that claim 32 stands rejected under 35 U.S.C. §103(a) as allegedly being unpatentable over U.S. Patent Number 2004/0093384 ("*Shipp*") in view of U.S. Publication Number 2004/0073617 ("*Milliken*") further in view of A Bayesian Approach to Filtering Junk E-Mail ("*Sahami*") further in view of Identifying Junk Electronic Mail In Microsoft Outlook with a Support Vector Machine ("*Woitaszek*"). Applicants respectfully traverse this rejection for at least the reason that *Shipp* further in view of *Milliken*, further in view of *Sahami*, further in view of *Woitaszek* fail to disclose, teach, or suggest all of the elements of claim 32. More specifically, claim 32 recites:

A computer-readable medium that includes a program that, when executed by a computer, performs at least the following:
receive an email message comprising an address, the email message further including displaying characters and non-displaying characters;

***search for the non-displaying characters in the email;
remove the searched non-displaying characters;***
tokenize the attachment to generate a token representative of the attachment;
determine a spam probability value from the generated token; and

sort the generated tokens in accordance with the corresponding determined spam probability value to determine a predefined number of interesting tokens, the predefined number of interesting tokens being a subset of the generated tokens,
wherein only displaying characters are tokenized.

(Emphasis added).

Applicants submit that claim 32, as amended, is allowable over the cited art for at least the reason that none of the references, taken alone or in combination, discloses, teaches, or suggests a "computer-readable medium that includes a program that, when executed by a computer, performs at least the following... ***search for the non-displaying characters in the***

email... [and] remove the searched non-displaying characters... wherein only displaying characters are tokenized' as recited in claim 32, as amended. While the Office Action argues that a combination of references renders claim 32 obvious, Applicants disagree. More specifically, *Shipp* discloses "[t]he invention is to weed out candidates for logging so that the normal mail is not logged. This reduces the burden on the database 23, and improves performance... A simplistic algorithm would be: If mail contains attachments, do not log" (page 3, paragraphs [0080] – [0081]). However, *Shipp* fails to even suggest a "computer-readable medium that includes a program that, when executed by a computer, performs at least the following... **search for the non-displaying characters in the email... [and] remove the searched non-displaying characters... wherein only displaying characters are tokenized'** as recited in claim 32, as amended.

Further, *Milliken* fails to overcome the deficiencies of *Shipp*. More specifically, *Milliken* discloses a "method for detecting transmission of potentially unwanted e-mail messages is provided. The method includes receiving e-mail messages and generating hash values based on one or more portions of the e-mail messages" (paragraph [0010]). However, *Milliken* fails to even suggest a "computer-readable medium that includes a program that, when executed by a computer, performs at least the following... **search for the non-displaying characters in the email... [and] remove the searched non-displaying characters... wherein only displaying characters are tokenized'** as recited in claim 32, as amended.

Additionally, *Sahami* fails to overcome the deficiencies of *Shipp* and *Milliken*. More specifically, *Sahami* discloses "*Sahami* discloses "[determining] whether a message has attached documents (most junk E-mail does not have them)... [is] also [a] powerful distinguisher between junk and legitimate E-mail" (page 3, right column, last paragraph). However, *Sahami* fails to even suggest a "computer-readable medium that includes a program that, when executed by a computer, performs at least the following... **search for the non-displaying**

characters in the email... [and] remove the searched non-displaying characters... wherein only displaying characters are tokenized" as recited in claim 32, as amended.

Further, *Woitaszek* fails to overcome the deficiencies of *Shipp*, *Milliken*, and *Sahami*. More specifically, *Woitaszek* discloses "[e]ach message [is] parsed to completely remove any headers, attachments, HTML markup, punctuation, and extended characters. This procedure essentially reduces a mail message to a series of delimited lowercase string tokens" (page 2, section 4). However, *Woitaszek* fails to even suggest a "computer-readable medium that includes a program that, when executed by a computer, performs at least the following... ***search for the non-displaying characters in the email... [and] remove the searched non-displaying characters... wherein only displaying characters are tokenized"*** as recited in claim 32, as amended. Thus, for at least the reason that none of the references disclose or suggest all of the claimed elements, the combination of references does not render obvious claim 32. For at least these reasons, claim 32, as amended, is allowable.

L. Claims 33 – 38 are Allowable Over *Shipp*, *Milliken*, *Sahami*, and *Woitaszek*

The Office Action indicates that claims 33 – 38 stand rejected under 35 U.S.C. §103(a) as allegedly being unpatentable over U.S. Patent Number 2004/0093384 ("*Shipp*") in view of U.S. Publication Number 2004/0073617 ("*Milliken*") further in view of A Bayesian Approach to Filtering Junk E-Mail ("*Sahami*") further in view of Identifying Junk Electronic Mail In Microsoft Outlook with a Support Vector Machine ("*Woitaszek*"). Applicants respectfully traverse this rejection for at least the reason that *Shipp* further in view of *Milliken*, further in view of *Sahami*, further in view of *Woitaszek* fail to disclose, teach, or suggest all of the elements of claims 32 – 38. More specifically, dependent claims 33 – 38 are believed to be allowable for at least the reason that these claims depend from and include the elements of allowable independent claim

32. *In re Fine, Minnesota Mining and Mfg. Co. v. Chemque, Inc.*, 303 F.3d 1294, 1299 (Fed. Cir. 2002).

CONCLUSION

In light of the foregoing amendments and for at least the reasons set forth above, all objections and/or rejections have been traversed, rendered moot, and/or addressed, and that the now pending claims are in condition for allowance. Favorable reconsideration and allowance of the present application and all pending claims are hereby courteously requested.

Any other statements in the Office Action that are not explicitly addressed herein are not intended to be admitted. In addition, any and all findings of inherency are traversed as not having been shown to be necessarily present. Furthermore, any and all findings of well-known art and Official Notice, or statements interpreted similarly, should not be considered well-known for the particular and specific reasons that the claimed combinations are too complex to support such conclusions and because the Office Action does not include specific findings predicated on sound technical and scientific reasoning to support such conclusions.

If, in the opinion of the Examiner, a telephonic conference would expedite the examination of this matter, the Examiner is invited to call the undersigned attorney at (770) 933-9500.

Respectfully submitted,

/afb/

Anthony F. Bonner Jr. Reg. No. 55,012

AT&T Legal Department – TKHR
Attn: Patent Docketing
One AT&T Way
Room 2A-207
Bedminster, NJ 07921
Customer No.: 38823